

reference expression level. In certain embodiments, the computer-implemented system displays a signal indicating that the subject has an increased likelihood of having emphysema. In certain embodiments, the computer-implemented system displays a signal indicating that the subject is in need of treatment for emphysema. In certain embodiments, the computer-implemented system displays a signal indicating the degree to which the expression levels in the sample obtained from a subject vary from those of the reference expression level. In certain embodiments, the computer-implemented system displays a signal indicating that the subject has an increased likelihood of having a more severe case of emphysema.

**[0023]** The details of various embodiments of the invention are set forth in the description below. Other features, objects, and advantages of the invention will be apparent from the description and the drawings, and from the claims.

#### DESCRIPTION OF THE DRAWINGS

**[0024]** FIG. 1 is an illustrated overview of the study design.

**[0025]** FIGS. 2A-2C depict maps of gene expression relevance networks. Dark blue circles indicate genes that have expression positively correlated with Lm while circles with white centers indicate all other genes. Edges are indicated by lighter (positive correlation) or darker (negative correlation) lines.

**[0026]** FIGS. 3A-3E shows the relationship between gene expression changes associated with regional emphysema severity (Lm) and cross-sectional studies of COPD-related gene expression using Gene Set Enrichment Analysis (GSEA). Black vertical lines represent the position of genes along the ranked gene list. The lengths of the black lines correspond to the magnitude of the running enrichment score from GSEA. Enrichments with an FDR q-value <0.05 were considered significant.

**[0027]** FIG. 3A shows that genes associated with Lm are enriched among the genes associated with emphysema or  $\alpha$ -1 antitrypsin deficiency (Goplon et al., American journal of respiratory cell and molecular biology 2004 31:595-600). T-statistic was calculated by a t-test between cases and controls. FIG. 3B shows that genes associated with Lm are enriched among the genes associated with diffusion capacity for carbon monoxide (DLCO) and forced expiratory volume at 1 sec (FEV1) (Spica et al American journal of respiratory cell and molecular biology 2004 31: 601-10), both of which are pulmonary phenotypic expressions of COPD. T-statistics were calculated by correlations of gene expression with a continuous variable. FIG. 3C shows that genes associated with Lm are enriched among the genes associated with DLCO, FEV1, FEV1/FVC (FVC=forced vital capacity), GOLD2 level emphysema, and GOLD3 level emphysema (Wang, L-M. et al. American journal of respiratory and critical care medicine 2008 177:402-11). T-statistics for DLCO, FEV1 and FEV1/FVC were calculated by Pearson correlation. T-statistics between COPD cases and never-smoked controls were calculated using a t-test. FIG. 3D shows that genes associated with Lm are enriched among the genes associated with FEV1/FVC, FEV1, and COPD diagnosis (Bhattacharya et al., American journal of respiratory cell and molecular biology 2009 40:359-67). T-statistics for FEV1 and FEV1/FVC were calculated by Pearson correlation. T-statistics between cases and controls were calculated using a t-test. FIG. 3E shows that genes previously found to be associated with COPD-related clinical variables are enriched

among the genes associated with Lm. T-statistic was calculated by correlations of gene expression with Lm using mixed-effect linear models.

**[0028]** FIG. 4 shows correlations demonstrating the relationship between gene expression changes associated with Lm and the cross-sectional stuffy of COPD-related gene expression from Goplon et al., (American journal of respiratory cell and molecular biology 2004 31:595-600). The graduated bar represents the t-statistic from a t-test between five emphysema patients and five non-smokers for 5,209 genes. The right end of the bar indicates a more positive t-statistic and left indicates a more negative t-statistic (induced or repressed in COPD, respectively). The vertical lines represent the position of genes associated with regional emphysema severity in the t-statistic ranking. The height of the vertical lines corresponds to the magnitude of the running enrichment score from GSEA.

**[0029]** FIG. 5 depicts correlations demonstrating the relationship between gene expression changes associated with Lm and those induced by TGF13 treatment of A549 cells (Malizia et al., American journal of physiology. Lung cellular and molecular physiology 2008 295:L451-60). The graduated bar represents the fold change between cell lines treated with and without TGF13 for 11910 genes. The vertical lines represent the position of genes associated with regional emphysema severity in the t-statistic ranking. The height of the vertical lines corresponds to the magnitude of the running enrichment score from GSEA.

**[0030]** FIGS. 6A-6D depict the relationship between gene expression changes associated with regional emphysema severity (Lm) and studies of TGF $\beta$ -related gene expression using GSEA. The vertical lines represent the position of genes associated with regional emphysema severity in the ranked gene list. The height of the vertical lines corresponds to the magnitude of the running enrichment score from GSEA. Enrichments with an FDR q-value <0.05 were considered significant.

**[0031]** FIG. 6A shows that genes associated with Lm are enriched among genes induced by TGF $\beta$  in Classen et al (Journal of immunology 2007 178:6931-40). FIG. 6B shows that genes associated with Lm are enriched among genes induced by TGF $\beta$  in Koinuma et al (Molecular and cellular biology 2009 29:172-86). FIG. 6C demonstrates that genes associated with Lm are enriched among genes induced by TGF $\beta$  in Malizia et al., (American journal of physiology. Lung cellular and molecular physiology 2008 295:L451-60). In FIGS. 6A-6C, the graduated bar represents the fold change between treated and untreated samples. FIG. 6D shows that genes most induced by TGF $\beta$  in seven studies are enriched among the genes that are associated with Lm. T-statistic was calculated by correlations of gene expression with Lm using mixed-effect linear models.

**[0032]** FIGS. 7A-7D depict the relationship between gene expression changes associated with regional emphysema severity (Lm) and the gene expression changes that occur when fibroblast cell lines are treated with GHK or TGF $\beta$ . The vertical lines represent the position of genes associated with regional emphysema severity in the ranked gene list. The height of the vertical lines corresponds to the magnitude of the running enrichment score from GSEA. Enrichments with an FDR q-value <0.05 were considered significant. FIG. 7A shows that genes whose expression levels increase in response to treatment with GHK or TGF $\beta$  are enriched among genes that decrease with increasing emphysema severity. The